CLAIMS

We claim:

1. A storage system, comprising:

one or more slow-access-time-mass-storage nodes, coupled to store data at respective first ranges of logical block addresses (LBAs);

a plurality of interim-fast-access-time nodes, configured to operate independently of one another, each interim-fast-access-time node being assigned a respective second range of the LBAs and coupled to receive data from and provide data to the one or more slow-access-time-mass-storage nodes having LBAs within the respective second range; and

one or more interface nodes, which are adapted to receive input/output (IO) requests from host processors directed to specified LBAs and to direct all the IO requests to the interim-fast-access-time node to which the specified LBAs are assigned.

2. A storage system according to claim 1, wherein the one or more interface nodes comprise a mapping between the interim-fast-access-time nodes and the LBAs, and wherein the one or more interface nodes are adapted to convert the IO requests to one or more requests and to direct the one or more requests to respective one or more interim-fast-access-time nodes in response to the mapping.

3. A storage system according to claim 2, wherein the mapping comprises a function relating each specific interim-fast-access-time node of the plurality of interim-fast-access-time nodes to the respective second range of the LBAs.

4. A storage system according to claim 2, wherein the

mapping comprises a table relating each specific interim-fast-access-time node of the plurality of interim-fast-access-time nodes to the respective second range of the LBAs.

5. A storage system according to claim 2, wherein the data is allocated into groups of data within the one or more slow-access-time-mass-storage nodes according to a pre-defined unit of the storage system comprising an integral number of bytes of the data, and wherein the mapping comprises a correspondence between the interim-fast-access-time nodes and the groups of data.

6. A storage system according to claim 1, wherein the one or more slow-access-time-mass-storage nodes comprise one or more disks, and wherein the interim-fast-access-time nodes comprise random access memories.

7. A storage system according to claim 1, wherein the plurality of interim-fast-access-time nodes comprise respective location tables, wherein each location table comprises locations of the second range of the LBAs assigned to the respective interim-fast-access-time node.

8. A storage system according to claim 1, wherein the respective second ranges are spread sufficiently evenly and finely so as to generate well-balanced loading for the plurality of interim-fast-access-time nodes.

9. A storage system according to claim 1, wherein each of the plurality of interim-fast-access-time nodes are at an equal hierarchical level.

10. A storage system according to claim 1, wherein the respective second ranges of the LBAs do not overlap.

11. A storage system according to claim 1, wherein the plurality of interim-fast-access-time nodes comprise a first and a second interim-fast-access-time node, and

wherein at least some of the respective second ranges of the LBAs of the first and the second interim-fast-access-time nodes comprise overlapping LBAs, so that one of the first and the second interim-fast-access-time nodes is operative as a redundant interim-fast-access-time node.

12. A storage system according to claim 1, wherein the one or more slow-access-time-mass-storage nodes comprise a multiplicity of slow-access-time-mass-storage nodes and the respective first ranges are spread sufficiently evenly and finely so as to generate well-balanced loading for the multiplicity.

13. A storage system according to claim 1, wherein the plurality of interim-fast-access-time nodes comprises a first interim-fast-access-time node and a second interim-fast-access-time node, and wherein the first and second interim-fast-access-time nodes have substantially equal capacities.

14. A storage system according to claim 1, wherein the plurality of interim-fast-access-time nodes comprises a first interim-fast-access-time node and a second interim-fast-access-time node, and wherein the first and second interim-fast-access-time nodes have different capacities.

15. A storage system according to claim 1, wherein the plurality of interim-fast-access-time nodes comprises a first interim-fast-access-time node and a second interim-fast-access-time node, and wherein the one or more slow-access-time-mass-storage nodes comprise a first slow-access-time-mass-storage node which is coupled to only receive data from and provide data to the first interim-fast-access-time node and a second slow-access-time-mass-storage node which is coupled to only receive data from and provide data to the second interim-fast-access-time node.

16. A storage system according to claim 1, wherein the plurality of interim-fast-access-time nodes comprises a first interim-fast-access-time node and a second interim-fast-access-time node, and wherein the one or more slow-access-time-mass-storage nodes comprise a first slow-access-time-mass-storage node and a second slow-access-time-mass-storage node which are coupled to receive data from and provide data to the first and the second interim-fast-access-time nodes.

17. A method for storing data, comprising:

storing the data in one or more slow-access-time-mass-storage nodes having respective first ranges of logical block addresses (LBAs);

assigning to each of a plurality of interim-fast-access-time nodes, configured to operate independently of one another, a respective second range of the LBAs;

coupling the plurality of interim-fast-access-time nodes to receive data from and provide data to the one or more slow-access-time-mass-storage nodes having LBAs within the respective second range;

receiving input/output (IO) requests from host processors directed to specified LBAs; and

directing all the IO requests to the interim-fast-access-time node to which the specified LBAs are assigned.

18. A method according to claim 17, wherein receiving the IO requests comprises providing one or more interface nodes, wherein the one or more interface nodes comprise a mapping between the interim-fast-access-time nodes and the LBAs, and wherein the one or more interface nodes are adapted to convert the IO requests to one or more LBA requests and to direct the one or more LBA requests to respective one or more interim-fast-access-time nodes in response to the mapping.

19. A method according to claim 18, wherein the mapping comprises a function relating each specific interim-fast-access-time node of the plurality of interim-fast-access-time nodes to the respective second range of the LBAs.

20. A method according to claim 18, wherein the mapping comprises a table relating each specific interim-fast-access-time node of the plurality of interim-fast-access-time nodes to the respective second range of the LBAs.

21. A method according to claim 18, wherein the data is allocated into groups of data within the one or more slow-access-time-mass-storage nodes according to a pre-defined unit comprising an integral number of bytes of the data, and wherein the mapping comprises a correspondence between the interim-fast-access-time nodes and the groups of data.

22. A method according to claim 17, wherein the one or more slow-access-time-mass-storage nodes comprise one or more disks, and wherein the interim-fast-access-time nodes comprise random access memories.

23. A method according to claim 17, wherein the plurality of interim-fast-access-time nodes comprise respective location tables, wherein each location table comprises locations of the second range of the LBAs assigned to the respective interim-fast-access-time node.

24. A method according to claim 17, wherein the respective second ranges are spread sufficiently evenly and finely so as to generate well-balanced loading for the plurality of interim-fast-access-time nodes.

25. A method according to claim 17, wherein each of the plurality of interim-fast-access-time nodes are at an equal hierarchical level.

26. A method according to claim 17, wherein the

respective second ranges of the LBAs do not overlap.

27. A method according to claim 17, wherein the plurality of interim-fast-access-time nodes comprise a first and a second interim-fast-access-time node, and wherein at least some of the respective second ranges of the LBAs of the first and the second interim-fast-access-time nodes comprise overlapping LBAs, so that one of the first and second interim-fast-access-time nodes is operative as a redundant interim-fast-access-time node.

28. A method according to claim 17, wherein the one or more slow-access-time-mass-storage nodes comprise a multiplicity of slow-access-time-mass-storage nodes and the respective first ranges are spread sufficiently evenly and finely so as to generate well-balanced loading for the multiplicity.

29. A method according to claim 17, wherein the plurality of interim-fast-access-time nodes comprises a first interim-fast-access-time node and a second interim-fast-access-time node, and wherein the first and second interim-fast-access-time nodes have substantially equal capacities.

30. A method according to claim 17, wherein the plurality of interim-fast-access-time nodes comprises a first interim-fast-access-time node and a second interim-fast-access-time node, and wherein the first and second interim-fast-access-time nodes have different capacities.

31. A method according to claim 17, wherein the plurality of interim-fast-access-time nodes comprises a first interim-fast-access-time node and a second interim-fast-access-time node, and wherein the one or more slow-access-time-mass-storage nodes comprise a first slow-access-time-mass-storage node which is coupled to only receive data from and provide data to the first interim-

fast-access-time node and a second slow-access-time-mass-storage node which is coupled to only receive data from and provide data to the second interim-fast-access-time node.

32. A method according to claim 17, wherein the plurality of interim-fast-access-time nodes comprises a first interim-fast-access-time node and a second interim-fast-access-time node, and wherein the one or more slow-access-time-mass-storage nodes comprise a first slow-access-time-mass-storage node and a second slow-access-time-mass-storage node which are coupled to receive data from and provide data to the first and the second interim-fast-access-time nodes.

33. A system for transferring data to and from one or more slow-access-time-mass-storage nodes which store data at respective first ranges of logical block addresses (LBAs), comprising:

a plurality of interim-fast-access-time nodes, configured to operate independently of one another, each interim-fast-access-time node being assigned a respective second range of the LBAs and coupled to receive data from and provide data to the one or more slow-access-time-mass-storage nodes within the respective second range; and

one or more interface nodes, which are adapted to receive input/output (IO) requests from host processors directed to specified LBAs and to direct all the IO requests to the interim-fast-access-time node to which the specified LBAs are assigned.

34. A system according to claim 33, wherein the one or more interface nodes comprise a mapping between the interim-fast-access-time nodes and the LBAs, and wherein the one or more interface nodes are adapted to convert

the IO requests to one or more requests and to direct the one or more requests to respective one or more interim-fast-access-time nodes in response to the mapping.

35. A system according to claim 34, wherein the mapping comprises a function relating each specific interim-fast-access-time node of the plurality of interim-fast-access-time nodes to the respective second range of the LBAs.

36. A system according to claim 34, wherein the mapping comprises a table relating each specific interim-fast-access-time node of the plurality of interim-fast-access-time nodes to the respective second range of the LBAs.

37. A system according to claim 34, wherein the data is allocated into groups of data within the one or more slow-access-time-mass-storage nodes according to a pre-defined unit of the storage system comprising an integral number of bytes of the data, and wherein the mapping comprises a correspondence between the interim-fast-access-time nodes and the groups of data.

38. A system according to claim 33, wherein the one or more slow-access-time-mass-storage nodes comprise one or more disks, and wherein the interim-fast-access-time nodes comprise random access memories.

39. A system according to claim 33, wherein the plurality of interim-fast-access-time nodes comprise respective location tables, wherein each location table comprises locations of the second range of the LBAs assigned to the respective interim-fast-access-time node.

40. A system according to claim 33, wherein the respective second ranges are spread sufficiently evenly and finely so as to generate well-balanced loading for the plurality of interim-fast-access-time nodes.

41. A system according to claim 33, wherein each of the

plurality of interim-fast-access-time nodes are at an equal hierarchical level.

42. A system according to claim 33, wherein the respective second ranges of the LBAs do not overlap.

43. A system according to claim 33, wherein the plurality of interim-fast-access-time nodes comprise a first and a second interim-fast-access-time node, and wherein at least some of the respective second ranges of the LBAs of the first and the second interim-fast-access-time nodes comprise overlapping LBAs, so that one of the first and the second interim-fast-access-time nodes is operative as a redundant interim-fast-access-time node.

44. A system according to claim 33, wherein the plurality of interim-fast-access-time nodes comprises a first interim-fast-access-time node and a second interim-fast-access-time node, and wherein the first and second interim-fast-access-time nodes have substantially equal capacities.

45. A system according to claim 33, wherein the plurality of interim-fast-access-time nodes comprises a first interim-fast-access-time node and a second interim-fast-access-time node, and wherein the first and second interim-fast-access-time nodes have different capacities.

46. A system according to claim 33, wherein the plurality of interim-fast-access-time nodes comprises a first interim-fast-access-time node and a second interim-fast-access-time node, and wherein the one or more slow-access-time-mass-storage nodes comprise a first slow-access-time-mass-storage node which is coupled to only receive data from and provide data to the first interim-fast-access-time node and a second slow-access-time-mass-storage node which is coupled to only receive data from and provide data to the second interim-fast-access-time

node.

47. A system according to claim 33, wherein the plurality of interim-fast-access-time nodes comprises a first interim-fast-access-time node and a second interim-fast-access-time node, and wherein the one or more slow-access-time-mass-storage nodes comprise a first slow-access-time-mass-storage node and a second slow-access-time-mass-storage node which are coupled to receive data from and provide data to the first and the second interim-fast-access-time nodes.

48. A method for transferring data to and from one or more slow-access-time-mass-storage nodes which store data at respective first ranges of logical block addresses (LBAs), comprising:

assigning to a plurality of interim-fast-access-time nodes, configured to operate independently of one another, respective second ranges of the LBAs;

coupling the plurality of interim-fast-access-time nodes to receive data from and provide data to the one or more slow-access-time-mass-storage nodes having LBAs within the respective second ranges;

receiving input/output (IO) requests from host processors directed to specified LBAs; and

directing all the IO requests to the interim-fast-access-time node to which the specified LBAs are assigned.

49. A method according to claim 48, wherein receiving the IO requests comprises providing one or more interface nodes, wherein the one or more interface nodes comprise a mapping between the interim-fast-access-time nodes and the LBAs, and wherein the one or more interface nodes are adapted to convert the IO requests to one or more LBA requests and to direct the one or more LBA requests to

respective one or more interim-fast-access-time nodes in response to the mapping.

50. A method according to claim 49, wherein the mapping comprises a function relating each specific interim-fast-access-time node of the plurality of interim-fast-access-time nodes to the respective second range of the LBAs.

51. A method according to claim 49, wherein the mapping comprises a table relating each specific interim-fast-access-time node of the plurality of interim-fast-access-time nodes to the respective second range of the LBAs.

52. A method according to claim 49, wherein the data is allocated into groups of data within the one or more slow-access-time-mass-storage nodes according to a pre-defined unit comprising an integral number of bytes of the data, and wherein the mapping comprises a correspondence between the interim-fast-access-time nodes and the groups of data.

53. A method according to claim 48, wherein the one or more slow-access-time-mass-storage nodes comprise one or more disks, and wherein the interim-fast-access-time nodes comprise random access memories.

54. A method according to claim 48, wherein the plurality of interim-fast-access-time nodes comprise respective location tables, wherein each location table comprises locations of the second range of the LBAs assigned to the respective interim-fast-access-time node.

55. A method according to claim 48, wherein the respective second ranges are spread sufficiently evenly and finely so as to generate well-balanced loading for the plurality of interim-fast-access-time nodes.

56. A method according to claim 48, wherein each of the plurality of interim-fast-access-time nodes are at an

equal hierarchical level.

57. A method according to claim 48, wherein the respective second ranges of the LBAs do not overlap.

58. A method according to claim 48, wherein the plurality of interim-fast-access-time nodes comprise a first and a second interim-fast-access-time node, and wherein at least some of the respective second ranges of the LBAs of the first and the second interim-fast-access-time nodes comprise overlapping LBAs, so that one of the first and second interim-fast-access-time nodes is operative as a redundant interim-fast-access-time node.

59. A method according to claim 48, wherein the plurality of interim-fast-access-time nodes comprises a first interim-fast-access-time node and a second interim-fast-access-time node, and wherein the first and second interim-fast-access-time nodes have substantially equal capacities.

60. A method according to claim 48, wherein the plurality of interim-fast-access-time nodes comprises a first interim-fast-access-time node and a second interim-fast-access-time node, and wherein the first and second interim-fast-access-time nodes have different capacities.

61. A method according to claim 48, wherein the plurality of interim-fast-access-time nodes comprises a first interim-fast-access-time node and a second interim-fast-access-time node, and wherein the one or more slow-access-time-mass-storage nodes comprise a first slow-access-time-mass-storage node which is coupled to only receive data from and provide data to the first interim-fast-access-time node and a second slow-access-time-mass-storage node which is coupled to only receive data from and provide data to the second interim-fast-access-time node.

62. A method according to claim 48, wherein the plurality of interim-fast-access-time nodes comprises a first interim-fast-access-time node and a second interim-fast-access-time node, and wherein the one or more slow-access-time-mass-storage nodes comprise a first slow-access-time-mass-storage node and a second slow-access-time-mass-storage node which are coupled to receive data from and provide data to the first and the second interim-fast-access-time nodes.